WHITE PAPER

A Plan for Cooperation Between NASA and DARPA to Establish a Center for Advanced Architectures

Peter J. Denning
George B. Adams III
Robert L. Brown
Pentti Kanerva
Barry M. Leiner
Michael R. Raugh

April 25, 1986

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report 86.10

RIACS

Research Institute for Advanced Computer Science

WHITE PAPER

A Plan for Cooperation Between NASA and DARPA to Establish a Center for Advanced Architectures

Peter J. Denning
George B. Adams III
Robert L. Brown
Pentti Kanerva
Barry M. Leiner
Michael R. Raugh

April 25, 1986

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report 86.10

TABLE OF CONTENTS

WHITE PAPER

A Plan for Cooperation Between NASA and DARPA to Establish a Center for Advanced Architectures

Peter J. Denning
George B. Adams III
Robert L. Brown
Pentti Kanerva
Barry M. Leiner
Michael R. Raugh

Research Institute for Advanced Computer Science
NASA Ames Research Center

RIACS Technical Report 86.10
April 25, 1986

Large, complex computer systems require many years of development. It is recognized that large-scale systems are unlikely to be delivered in useful condition unless users are intimately involved throughout the design process.

This report describes a mechanism that will involve users in the design of advanced computing systems and will accelerate the insertion of new systems into scientific research. This mechanism is embodied in a facility called the *Center for Advanced Architectures* (CAA). CAA would be a division of RIACS and would receive its technical direction from a Scientific Advisory Board established by RIACS.

The CAA described here is a possible implementation of a center envisaged in a proposed cooperation between NASA and DARPA.

Work reported herein was supported in part by
Contract NAS2-11530 and by Cooperative Agreement NCC 2-387
from the National Aeronautics and Space Administration (NASA)
to the Universities Space Research Association (USRA).

WHITE PAPER

A Plan for Cooperation Between NASA and DARPA to Establish a Center for Advanced Architectures

Peter J. Denning
George B. Adams III
Robert L. Brown
Pentti Kanerva
Barry M. Leiner
Michael R. Raugh

Research Institute for Advanced Computer Science

April 25, 1986

1. PROBLEM STATEMENT

Large, complex computer systems require many years of development. It is recognized that large-scale systems are unlikely to be delivered in useful condition unless users are intimately involved throughout the design process.

This report describes a mechanism that will involve users in the design of advanced computing systems and will accelerate the insertion of new systems into scientific research. This mechanism is embodied in a facility called the *Center for Advanced Architectures* (CAA). CAA would be a division of RIACS and would receive its technical direction from a Scientific Advisory Board established by RIACS.

The CAA described here is a possible implementation of the center envisaged in a proposed cooperation between NASA and DARPA.

2. GENERAL PURPOSES

A. The CAA will provide a national focus for exploring applicability of new computer architectures in scientific research.

a. "Exploring" means conducting research leading to deep understanding of how particular architectures can match the needs of particular scientific domains. A few explorations will be conducted by RIACS, but most will be conducted by outside groups using administrative mechanisms provided by RIACS.

b. All explorations are intended to be in depth, meaning they take up intimate details of the architectures and the domains. Explorations will emphasize interactions between scientist-users and architects; and between software and hardware of systems.

c. The domains of primary interest are scientific. Examples are computational fluid dynamics, computational chemistry, flight simulation, galactic modeling, search for extra terrestrial intelligence, space automation and robotics, image processing, cognitive modeling, and speech and text recognition.

d. The technologies to be explored include hardware, especially parallel machines and special-purpose machines; communications, especially very-high-speed networks, multi-media conferencing, and internets; and software, especially programming environments and new algorithms for solving problems in specific domains.

B. Explorations will intimately involve potential end-users (working scientists in the domains) with hardware and software architects. Some combinations of domain and architecture will doubtless be judged unmatched, while others will be judged good matches. Actual experience using the test systems is intended to influence the design. Results of these explorations will be architectures better matched to the domains, software that aids programming, new algorithms for key problems, and new techniques for solving problems in the domains.

C. The purpose of these explorations is to accelerate the integration of new computing technologies into scientific research processes. This will in turn accelerate, and expand the scope of, scientific research in the U.S.

D. In the long term, industry, government, and academia will be partners in the CAA.

a. Industry will be the primary source of new architectures. Government and Academia will be the primary sources of scientists in the various disciplines. Government would also be a source of support for specific research projects to test new architectures. RIACS would not attempt to fund or sponsor all explorations that may be proposed.

b. RIACS will provide a mechanism to review proposals for new machines before they are incorporated into the CAA. At a minimum, a new machine will have to be reliable, have a credible programming environment, and be of interest in several disciplines.

c. All partners (industry, academia, government) will be strongly encouraged to place personnel on site through a visiting scientist/architect program. On-site visiting architects will be an important means for passing discussion-group experience back to the manufacturer.

d. The availability of the CAA, support of exploration groups, and access to new technologies will be an incentive for scientists to participate in explorations. Because they provide a means for exercising new architectures in scientific disciplines, these same mechanisms will be an incentive for manufacturers to place early prototypes in the testbed and to match new systems to groups of users.

E. The CAA will pursue high standards of quality: "If RIACS installs a machine in the CAA testbed, that machine is worth looking at." This implies careful evaluation of machines before they are installed. It implies involvement of top scientists and architects in policy-making and selection of architectures.

3. MECHANISM

A. RIACS will set up a new division to administer CAA. The CAA will contain the testbed systems, support and adminstrative staff, tool developers, and access to the National Research Internet (NRI). The CAA will not be a general-purpose, production-oriented computing center; it will serve a selected community of researchers participating in experiments to use and explore the capabilities of the machines.

a. It is not necessary that all machines of the CAA be physically present at the facility. They can be sited elsewhere, e.g., at the manufacturer's laboratory, but be accessible via the NRI. The CAA will organize all administrative interactions with the scientist users, so that the location of a machine is invisible.

b. It is essential that the CAA be part of the NRI so that it is accessible to the entire scientific community of the U.S. This suggests that the CAA will employ DARPA standards for internetworking, distributed computing, multimedia conferencing, etc.

c. It is essential that the administration of CAA not interfere with RIACS's charter to conduct research. The CAA will be managed by an experienced professional, who reports to the RIACS Director. This arrangement will strengthen both the RIACS research program and the scientific value of the CAA.

B. RIACS will promote and facilitate *discussion groups* that explore specific combinations of architecture and domain. The purpose of discussion groups is to gain and share experience in using new architectures and make this information available to the machine's architects and to the community.

a. A discussion group is a collection of scientists, architects, and programmers who are studying the feasibility of an architecture in selected domains. Participants should actively attempt to program the machine, then report their experiences to the rest of the group.

b. Active participation in a discussion group should be a prerequisite for access to one of the machines in the CAA.

c. Some discussion groups will explore the utility of an architecture in a given domain. Others will take up issues that span many domains, such as theory of parallel computing.

d. Discussion groups can take a variety of forms. Examples: biweekly groups of CFD (computational fluid dyanmics) researchers at Ames Research Center exploring hypercube architectures; a group of researchers convened for a two-week workshop; a group organized and operating over the network without face-to-face meetings.

e. The CAA will develop mechanisms to encourage formation of groups, to assist them conduct their business, to obtain reports on their findings, and make the results available to the NRI community.

C. RIACS will conduct its own program of research in selected domain-architecture combinations. This will ensure that RIACS scientists and CAA staff understand first-hand the processes being used in the discussion groups. It will also help attract top scientists to RIACS on visiting appointments to participate more intimately in the evaluation of new architectures.

a. Such research is underway already at RIACS. Examples include: study of MIT Dataflow Machine; feasibility of CFD algorithms on the Intel hypercube, conducted by John Bruno of UC Santa Barbara; and an Ames/Stanford discussion group led by Tony Chan on CFD algorithms on the hypercube.

b. The RIACS research program will be expanded. Possibilities now under discussion at Ames are: use of hypercube in computational chemistry; parallel machines for flight simulation; and connection machine for fluid flow

simulation and cognitive modeling.

c. The RIACS studies will be a mixture of bottom-up and top-down. Bottom-up studies will consider the feasibility of using a specific machine in various domains. Top down studies will consider problems in specific domains and determine architectural characteristics best suited for those problems.

D. RIACS will conduct its own program of research in computer architecture. This will ensure that RIACS scientists and CAA staff understand first-hand the process of incorporating domain information into an architecture.

a. Pentti Kanerva's Sparse Distributed Memory (SDM) project is already underway and will serve this purpose. Discussion groups can be formed in areas of memory simulation, robot learning, text and speech recognition, vision, and system interfacing.

E. RIACS will organize a *Scientific Advisory Board* (SAB) consisting of scientists from government, academia, and industry to help set general policy directions for the CAA and receive input from various communities. In addition to the SAB, RIACS will use specific selection committees will evaluate possible new additions to the testbed.

a. RIACS's parent organization, USRA (Universities Space Research Association), has considerable experience in organizing science councils made of leading scientists. They have contacts in the 58 member universities to draw on.

b. Selection committees will be formed to evaluate proposed new additions to the CAA. Proposals for new additions will be solicited from the entire community of manufacturers and researchers.

c. The SAB will evaluate scientific progress of the various discussion groups and will help decide when groups have completed their work and when machines have completed their usefulness in the facility.

4. WHY RIACS?

A. RIACS is operated by USRA, a consortium of 58 universities with many contacts throughout the scientific community.

B. RIACS is in a unique position to help conduct these studies because it is in direct, active contact with scientists at NASA Ames Research Center, at several universities (Stanford, UC Berkeley, UC Santa Barbara, UC San Diego, Caltech,

Duke, Yale, and U Toronto), and at several computer manufacturers (Sequent, Intel, and Sun).

C. RIACS has experience in the multi-disciplinary studies and evaluation groups of the type envisaged here.

a. In 1984, we conducted a study of the MIT static dataflow machine, which provided detailed information supportive of the commercial feasibility of a dataflow machine. We issued a report on this study and a general report on how to conduct studies of this type.

b. In 1984, we conducted a feasibility study of the MPP (Massively Parallel Processor) for CFD problems.

c. In 1985, we conducted a feasibility study of the Intel hypercube for CFD problems.

d. In 1986, we are running a discussion group with CFD and Stanford researchers and the hypercube.

D. RIACS already operates a parallel machine testbed. It consists of a Sequent Balance 8000 and an Intel iPSC hypercube; these machines are accessible on the Ames network and on the DARPA internet. Discussion groups are underway concerning these machines and CFD.

E. RIACS has projects leading to programming environments for parallel machines -- Concurrent C and graphical shell. These tools will be used in our own studies and will be useful in other systems.

F. RIACS has projects leading to new parallel algorithms on shared-memory and hypercube architectures for problems in computational fluid dynamics (Tony Chan, John Bruno), computational chemistry (Harry Partridge), and image processing (George Adams).

G. RIACS is situated squarely in the middle of a rich computational environment at the NASA Ames Research Center. Facilities to which RIACS has access include: multi-architecture supercomputers in the Ames Computing Facility; supercomputers, notably the Cray 2, and other test machines in the NAS (Numerical Aerodynamic Simulator) facility, which is a NASA pathfinder in computational resources for aerospace research; gateways to the national research internet (NRI); connections to ARPANET, NSFNET, CSNET, NASA PCSN, University of California network, and other networks used in the scientific community.

5. IMPLEMENTATION PLAN

A. We propose a two-year startup project to be evaluated by NASA and DARPA after 18 months. If the startup project is successful, RIACS would submit additional proposals to continue and expand the facility.

B. The initial CAA facility would consist of a 64K-node Connection Machine, a front-end host, several workstations, and access to the NRI.

C. The CAA would be administratively a separate division of RIACS, consisting initially of these personnel:

a. A professional manager of the CAA, who would report to, and receive scientific direction from, the RIACS Director. The CAA manager would have full responsibility for the operation and administration of the CAA. The remaining CAA personnel would report to the CAA manager.

b. A systems programmer, who would be responsible for the operating systems software of the CAA, programming environments, libaries, and maintenance. (Note: If the first architecture in the facility is a Connection Machine, the systems programmer would be provided as part of the maintenance agreement with Thinking Machines Corporation. This staff position would remain vacant until additional machines were added to the CAA.)

c. A network programmer, who would be responsible for access of CAA machines to national networks, and for installing and developing network tools to support discussion groups.

d. An applications programmer to assist the two initial discussion groups carry out experiments on the machines.

e. A secretary to handle the public interface, inquiries, mailings, information, and the like.

In addition to the above CAA staff, RIACS would need to add an administrative assistant to support the scientific activities of the discussion groups.

We believe it is essential that the CAA provide the basic structure for supporting discussion groups. The staff above are sufficient to do this during the initial 18-month period. Agencies who support additional discussion groups would be expected to contribute a fair share for the CAA staff that support those new groups.

D. RIACS would organize the Scientific Advisory Board (SAB) and hold the first meeting to determine initial directions of the CAA. The SAB would meet twice annually.

E. RIACS would carry out research on a new, massively parallel architecture for pattern-oriented computing, Kanerva's *Sparse Distributed Memory* (SDM). A simulator for SDM would be implemented on the Connection Machine and would be used in an SDM discussion group. Experience with applications would guide architectural improvements of the simulator, which would be reflected in the hardware SDM being designed in a companion project with Michael Flynn at Stanford University. (Descriptions of the SDM project have been given separately.)

F. We intend that this project sponsor a small number (2 or 3) of discussion groups based on the Connection Machine.

a. Applications of Sparse Distributed Memory.

b. Application of the connection machine architecture for fluid dynamics problems, especially turbulence modeling. (Some work has already begun at other sites in this area.)

To assure scientific leadership of these groups, RIACS would request support for two additional visiting scientists.

6. NASA and DARPA Cooperation in the CAA

A. NASA and DARPA can contribute unique resources to the CAA and thereby accomplish something that neither agency could do alone. NASA includes a large, scientific research community spanning a wide variety of scientific disciplines, many of which seek access to the computing power of new, parallel and distributed computing architectures. DARPA has been sponsoring rapid development of new architectures and communication technologies; they seek testing and deployment of the best of these technologies in the scientific workplace.

The program outlined below consists of establishment of the CAA and mechanisms to make effective use of that facility. In general, DARPA will use the CAA to test DARPA-sponsored architectures with respect to satisfying scientific computing requirements. NASA will encourage its scientific community to test the capabilities of new architectures in the CAA with respect to NASA research domains.

B. We propose that NASA and DARPA sign a Memorandum of Understanding (MoU) to establish the CAA described in this report. The period of the MoU would be five years, with annual review by the two agencies beginning after

18 months. The two agencies would periodically specify the resources each intends to contribute to the project for the next stage.

C. We propose that the initial stage focus on establishing the facility with a single, broadly applicable machine, two discussion groups in RIACS centered on this machine, and an architecture project in RIACS that would use this machine as a simulator during the design. This initial stage would be complete in 12-18 months. Additional stages would expand the the facility after sufficient experience is gained to ensure smooth operation.

D. For the initial stage, NASA might provide these resources:

a. Fully serviced machine room space.

b. Wide-band connection to National Research Internet (NRI).

c. Offices for CAA staff.

d. Support for the Sparese Distributed Memory (SDM) project as the initial architecture research project envisaged in RIACS as outlined earlier.

e. Support for at least two discussion groups of interest to NASA. Potential areas are flight simulation, fluid dynamics, chemistry, neural modeling, and autonomous systems.

In addition, NASA would provide the following as part of its general program of research that makes possible the effective use of the CAA:

f. Access through NAS (Numerical Aerodynamic Simulator) Projects Office to test machines situated in NAS facilities;

g. Support for visiting-scientist feasibility studies of specific architecture-domain combinations.

E. For the initial phase, DARPA might provide the these resources:

a. 64K node Connection Machine and associated equipment. Additional machines would be provided as the center evolves and the concept validated.

b. Support of the staff to run the facility itself (manager, programmers, technicians, secretary).

c. Maintenance of CAA equipment.

d. Support for visiting scientists.

e. Administrative personnal for support of discussion groups and publications.

In addition, DARPA would provide the following as part of its general program of research:

f. Access to relevant DARPA software such as multi-media conferencing and distributed OS.

g. Use of the Internet communication system including but not limited to the ARPANET and Wideband Satellite Network.